Journal of Network and Computer Applications 64 (2016) 62–75

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Smartphone-based crowdsourcing for estimating the bottleneck capacity in wireless networks

Enrico Gregori [a], Alessandro Improta [a], Luciano Lenzini [b], Valerio Luconi [b], Nilo Redini [b], Alessio Vecchio [b,*]

[a] IIT-CNR, Pisa, Italy
[b] Dip. di Ingegneria dell'Informazione, Università di Pisa, Pisa, Italy

ARTICLE INFO

Article history:
Received 4 February 2015
Received in revised form
27 November 2015
Accepted 19 January 2016
Available online 6 February 2016

Keywords:
Crowdsourcing
Smartphones
Bottleneck capacity
Network tools

ABSTRACT

Crowdsourcing enables the fine-grained characterization and performance evaluation of today's large-scale networks using the power of the masses and distributed intelligence. This paper presents SmartProbe, a system that assesses the bottleneck capacity of Internet paths using smartphones, from a mobile crowdsourcing perspective. With SmartProbe measurement activities are more bandwidth efficient compared to similar systems, and a larger number of users can be supported. An application based on SmartProbe is also presented: georeferenced measurements are mapped and used to compare the performance of mobile broadband operators in wide areas. Results from one year of operation are included.

1. Introduction

Crowdsourcing exploits the help of the masses to solve large-scale problems. A task that is too demanding for the internal resources of a single organization can be divided into small and loosely-coupled activities which are assigned to and carried out by a population of individuals (Howe, 2008). Unlike outsourcing, with crowdsourcing the identity of the cooperating users is generally not relevant, and the workforce dynamically changes according to the necessities of the delegating organization and the will of the participants. Crowdsourcing is used in a variety of contexts, from the production of creative content to the massive analysis of data. Notable examples include Threadless (2015), an online clothes shop where the design of items is collaborative and user-driven, and SETI@home, a distributed effort aimed at searching for extraterrestrial intelligence using spare CPU cycles of users' machines (SETI, 2015). There are also several platforms that support crowdsourcing-based interaction, such as Mechanical Turk (Amazon, 2015), Microworkers (2015), and Crowdflower (2015). These platforms provide other companies with methods for submitting tasks, enrolling users, and managing payments. Although in some cases the activities delegated to users are trivial, in other situations the tasks are complex and require intelligence and/or creativeness. In all cases, the whole result is greater than the sum of its parts, and decentralized and distributed intelligence, aggregated through crowdsourcing, can provide an answer to unsolved scientific and engineering problems.

Crowdsourcing systems have traditionally been based on the web, as it provides collaboration tools that are both efficient and easy to use (Doan et al., 2011). More recently, the web-centric interaction model has been expanded to support smartphone-based cooperation (Chatzimilioudis et al., 2012). In fact, smartphones are an appealing platform for crowdsourcing applications: they are always on and carried by their owners, they are mobile and richly connected, and they are equipped with an increasing number of sensors (camera, microphone, etc). When using a smartphone, the working user is no longer constrained to a fixed position and he/she can carry out the requested task at different locations, possibly using additional input mechanisms. The term *crowdsensing* is used when sensing is the prevalent activity delegated to participants. Crowdsensing applications can be classified according to the type of phenomenon being measured. Examples include environmental applications (for observing pollution and water levels), infrastructure monitoring applications (for collecting information about traffic congestions and road conditions), and social applications (for monitoring activity levels of individuals) (Ganti et al., 2011; Vivacqua and Borges, 2012). In several scenarios, the crowdsensing activities are carried out without a well-defined employer-employee relationship. Users may be interested in participating in sensing activities for a number of reasons, including altruism or the scientific relevance of the end goals. In other cases they perceive the results of the sensing activity as also

* Correspondence to: Dip. di Ingegneria dell'Informazione, Università di Pisa, Largo L. Lazzarino 1, 56122 Pisa, Italy.

http://dx.doi.org/10.1016/j.jnca.2016.01.020
1084-8045/

being useful for themselves (even though the small task each user completes may be scarcely significant, the level of interest in the aggregated results can be much higher).

An increasing number of crowdsourcing/crowdsensing systems are related to networking: the crowd-based approach provides a solution to the need for collecting detailed information on today's large-scale networked systems. For instance crowdsourcing is currently used for the following purposes: (*i*) detection of traffic differentiation silently applied by Internet service providers to their customers (Dischinger et al., 2010), (*ii*) characterization of the Internet and detection of network problems (Shavitt and Shir, 2005; Sánchez et al., 2013), (*iii*) analysis and measurement of wireless networks (Rai et al., 2012).

This paper presents SmartProbe, a network measuring tool designed to operate in a mobile crowdsensing scenario. Using smartphones as measuring elements, SmartProbe estimates the bottleneck capacity of Internet paths. Since it is executed on user devices, SmartProbe was designed and customized to generate less traffic, and thus to use less energy, compared to similar systems (experimental results show a significant reduction in sent/received data). In addition, since the system has to be used by a possibly large number of users, the software infrastructure on the server-side was designed to cope with multiple measurement requests. Measurements can be georeferenced thanks to the self-positioning ability of smartphones. A demo application is also presented: different mobile broadband operators are compared using crowdsourced measurements. Other possible uses include mapping the performance of Wi-Fi access points in an urban area or analyzing the performance of a cell phone operator in relation to variations in user positions.

2. Background and motivation

Let L_i be the i-th link of an Internet path, with $i \in 1 .. n$. The capacity C_i of the i-th link is the maximum data rate that link L_i can achieve. The bottleneck capacity of an Internet path is defined as the capacity of the narrowest link of the path considered (Dovrolis et al., 2003), and thus it is equal to min $\{C_1, \ldots, C_n\}$ (in other words it is the capacity of the link that imposes a bottleneck on the path in terms of data rate). The capacity of a link (or a path) is a static property that does not change with time, and should not be confused with the available bandwidth. The latter represents the residual bandwidth not currently in use by other traffic (a property whose value depends on current conditions) (Jain et al., 2004; Jain and Dovrolis, 2003). If $A_i(t)$ is the available bandwidth of L_i at time t, then $C_i \geq A_i(t)$ holds. For example, the bottleneck capacity of the path shown in Fig. 1 is C_3, the capacity of L_3, whereas the link with the smallest available bandwidth is L_5 (considering the current utilization level as depicted by the gray areas).

The ability to measure the bottleneck capacity is useful not only for network protocols (e.g. for congestion control) but also at the application and user levels. In fact, this information can be used to tune the operational parameters of streaming applications and peer-to-peer systems, or to evaluate the actual performance of a residential ISP connection.

Methods and techniques for measuring this network property have received significant attention from both researchers and practitioners (Dovrolis et al., 2004; Kapoor et al., 2004; Chen et al., 2008). In addition, while most of the initial activities have been carried out for wired networks, more recently the study of techniques specifically designed for wireless environments has gained momentum (Chen et al., 2009; Li et al., 2003). In this paper we move forwards in two different directions. On the one hand, we continue this trend by giving even more relevance to wireless scenarios. The aim is to make the evaluation of the bottleneck capacity more suitable for execution on mobile devices (smartphones and tablets, which have surpassed common PCs in terms of sales, now represent the preferred Internet-enabled devices for the majority of users). On the other hand, we believe that an evolution of bottleneck capacity estimation tools in a crowdsourcing perspective may pave the way for interesting and unexplored usages, as it may provide fine-grained information on today's large scale networked systems.

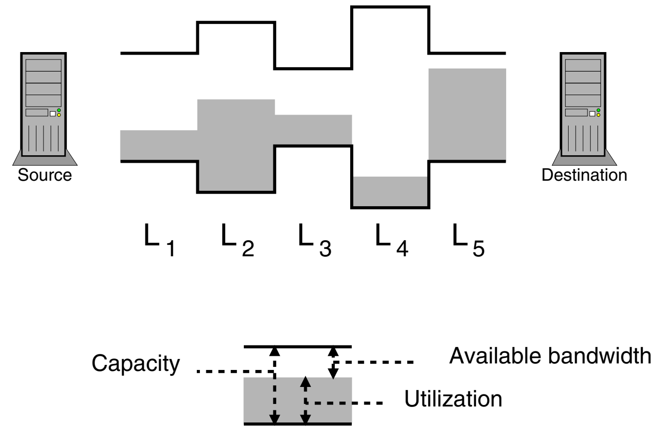

Fig. 1. An Internet path with links characterized by different capacities and utilization levels.

The evaluation of the bottleneck capacity in a smartphone-based crowdsourcing scenario must take into account (*i*) bandwidth and energy efficiency, (*ii*) tuning for wireless connections, and (*iii*) support for possibly large numbers of users.

Bandwidth and energy efficiency. On smartphones and tablets, energy is a scarce resource and communication can be expensive. As a consequence, tools that do not consider these factors are likely to be abandoned by their users. Almost all the techniques designed and implemented so far have been conceived with the assumption that the devices are connected via wired links. Thus, they are not particularly efficient from this point of view.

In SmartProbe, the estimation of the bottleneck capacity has been designed to be suitable for mobile devices, with bandwidth efficiency as a primary goal. Reducing the amount of data transmitted and received, in turn, brings advantages in terms of energy efficiency. Our techniques considerably reduce the amount of traffic: up to 96% for 3G cellular networks and up to 89% for Wi-Fi networks.

Tuning for wireless connections. Most existing techniques do not behave properly in wireless scenarios, which tend to have high bit error rates. In addition, the tools currently available for smartphones implement rather trivial techniques: they perform an HTTP GET request and calculate the time needed to retrieve the requested page. However what is obtained is just an estimation of *TCP throughput* and not of bottleneck capacity. These are actually two different properties that should not be confused (Jain and Dovrolis, 2003) (for instance, TCP throughput depends on a number of factors such as buffer size, round trip time, and retransmission errors).

SmartProbe introduces customizations to state-of-the-art techniques in order to operate more smoothly in wireless scenarios.

Support for possibly large numbers of users. The use of bottleneck estimation methods from a crowdsourcing perspective creates a new set of problems. Countermeasures are needed to prevent simultaneous measurements from interfering with each other.